

08CV3742

JUDGE PALLMEYER

MAGISTRATE JUDGE VALDEZ

TG

EXHIBIT B



US006518414B1

(12) **United States Patent**
MacLennan

(10) **Patent No.:** **US 6,518,414 B1**
(45) **Date of Patent:** **Feb. 11, 2003**

(54) **MOLECULAR CLONING AND EXPRESSION OF G-PROTEIN COUPLED RECEPTORS**

(76) **Inventor:** **Alexander John MacLennan**, 7811 NW 35th Place, Gainesville, FL (US) 32606

(*) **Notice:** Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) **Appl. No.:** **09/225,024**

(22) **Filed:** **Jan. 4, 1999**

Related U.S. Application Data

(63) Continuation-in-part of application No. 08/760,936, filed on Dec. 6, 1996, now Pat. No. 5,856,443, which is a continuation of application No. 08/196,989, filed on Feb. 15, 1994, now Pat. No. 5,585,476.

(51) **Int. Cl.⁷** **C12N 15/12**
(52) **U.S. Cl.** **536/23.5; 435/69.1**
(58) **Field of Search** **435/7.1, 7.2; 475/69.1; 530/350; 536/23.5**

(56) **References Cited**

PUBLICATIONS

An et al. Identification of cDNAs encoding two G protein-coupled receptors for lysosphingolipids. *FEBS Letters* 417:279-282, 1997.*
Jackson. Structure and Function of G Protein Coupled Receptors. *Pharmac. Ther.* 50:425-442, 1991.*
Supplemental Information Disclosure Statement Under 37 CFR §§1.97 and 1.98 dated Mar. 14, 2002.
Declaration of Stephen Sims dated Nov. 8, 2001.
Yarden, Y.A. Ullrich (1988) "Growth Factor Receptor Tyrosine Kinases" *Ann. Rev. Biochem.* 57:443-478.
Devreotes, P. (1989) "*Dictyostelium discoideum*: A Model System for Cell-Cell Interactions in Development" *Science* 245:1054-1058.
Hanley, M.R. (1989) "Mitogenic neurotransmitters" *Nature* 340:97.
Zachary, I., P.J. Woll, E. Rozengurt (1987) "A Role for Neuropeptides in the Control of Cell Proliferation" *Dev. Biol.* 124:295-308.
Young, D., G. Waitches, c. Birchmeier, O. Fasano, M. Wigler (1986) "Isolation and Characterization of a New Cellular Oncogene Encoding a Protein with Multiple Potential Transmembrane Domains" *Cell* 45:711-719.

Gutkind, J.S., E.A. Novotny, M.R. Brann, K.C. Robbins (1991) "Muscarinic acetylcholine receptor subtypes as agonist-dependent oncogenes" *Proc. Natl. Acad. Sci. USA* 88:4703-4707.

Julius, D., T.J. Livelli, T.M. Jessell, R. Axel (1989) "Ectopic Expression of the Serotonin 1c Receptor and the Triggering of Malignant Transformation" *Science* 244:1057-1062.

Julius, D., K.N. Huang, T.J. Livelli, R. Axel, T.M. Jessell (1990) "The 5HT₂ receptor defines a family of structurally distinct but functionally conserved serotonin receptors" *Proc. Natl. Acad. Sci. USA* 87:928-932.

MacLennan, A.J., G.D. Frantz, R.C. Weatherwax, N.J.K. Tillakaratne, A.J. Tobin (1990) "Expression of mRNAs That Encode D₂ Dopamine Receptor Subtypes: Anatomical, Developmental, and Pharmacological Studies" *Molec. Cell. Neurosci.* 1:151-160.

Loy, E.Y., J.F. Elliott, S. Cwirla, L.L. Lanier, M.M. Davis (1989) "Polymerase Chain Reaction with Single-Sided Specificity: Analysis of T Cell Receptor α Chain" *Science* 243:217-220.

Sanger, F., S. Nicklen, A.R. Coulson (1977) "DNA sequencing with chain-terminating inhibitors" *Proc. Natl. Acad. Sci. USA* 74:5463-5467.

Chirgwin, J.M., E. Przbyla, R.J. MacDonald, W.J. Rutter (1979) "Isolation of Biologically Active Ribonucleic Acid from Sources Enriched in Ribonuclease" *Biochem.* 18:5294-5299.

Okazaki (1993) "Molecular Cloning of a Novel Putative G Protein Coupled Receptor Expressed in the Cardiovascular System" *Biochem. Biophys. Res. Comm.* 190:1104-1109.

* cited by examiner

Primary Examiner—John Ulm

(74) *Attorney, Agent, or Firm*—Saliwanchik, Lloyd & Saliwanchik

(57) **ABSTRACT**

The subject invention pertains to cloning and expression of novel cDNAs which encode members of the G-protein coupled receptor superfamily of proteins. Polynucleotides which encode mammalian H218 protein are described. The invention also concerns methods for screening for ligands of H218 protein. The proteins and peptides of the subject invention can also be used to produce antibodies which can bind to the subject proteins. The polynucleotide molecules, proteins, and antibodies of the subject invention can be used in both diagnostic and therapeutic applications.

2 Claims, 13 Drawing Sheets

U.S. Patent

Feb. 11, 2003

Sheet 1 of 13

US 6,518,414 B1

-155 CCCCCCCCCCTCGAGCACAGCCACAGTACCAAGTCAGCCACTGGTGTCCCGG
-95 GCGCAGACGCCAAGCCACTCAGGCCAGGACCCCTGGCCGCTAGCCAGTGCT
-35 CAGTCCCATGGCCCCCGGCCACTGAGCCGCGGCGGTTTATCTACTCAGAGTAC
MetGlyGlyLeuTyrSerGlyTyr 8
25 CTCAATCCTGAGAAGGTTCAAGAACACTACATAACCAAGGAGACGCTGGACATGCAG
LeuAsnProGluLysValGlnGluHisTyrAsnTyrThrLysGluThrLeuAspMetGln 28
85 GAGACGCCCTCCGCAAGGTGGCCTCCGCTTCATCATCATTTTATGTGTGCCATCGTG
GluThrProSerArgLysValAlaSerAlaPheIleIleLeuCysCysAlaIleVal 48
145 GTGGAGAACCTTCTGGTCTAATCGCAGTGGCCAGGAACAGCAAGTTCCTCAGCCATG
ValGluAsnLeuLeuValIleAlaValAlaArgAsnSerLysPheHisSerAlaMet 68
205 TACCTGTTCTCGGCAACCTGGCAGCCTCCGACCTGCTGGCAGCGGTGGCTTCGTGGCC
TyrLeuPheLeuGlyAsnLeuAlaAlaSerAspLeuAlaGlyValAlaPheValAla 88
265 AACACCTTGCTCTCGGACCTGTCAACCTGTCTTAACCTTGCAGTGGTTGCCCGA
AsnThrLeuLeuSerGlyProValThrLeuSerLeuThrProLeuGlnTrpPheAlaArg 108
325 GAGGTTACGCTTCATCAGCTCTCTGCTCGCTTTCAGCCCTCCTGGCCATTGCCATC
GluGlySerAlaPheIleThrLeuSerAlaSerValPheSerLeuLeuAlaIleAlaIle 128
385 GAGAGACAAGTGGCCATCGCCCAAGGTCAAGCTCTACGGCAGTCACAAAAGCTGTGCAATG
GluArgGlnValAlaIleAlaLysValLysLeuTyrGlySerAspLysSerCysArgMet 148
445 TTGATCCTCATTTGGGCTCTTGGCTGATATCGCTGATTCTGGTGGCTTGGCCATCCTG
LeuMetLeuIleGlyAlaSerTrpLeuIleSerLeuIleLeuGlyGlyLeuProIleLeu 168
505 GCCTGGAATTGCTGGACCATCTGGAGGCTTGTCCACTGTGTGCCCTCTATGCTAAG
GlyTrpAsnCysLeuAspHisLeuGluAlaCysSerThrValLeuProLeuTyrAlaLys 188
565 CACTATGTGCTCTGCGTGGTCAACCATCTTCTGTCTACTGTGCTATCGTGGCCTTG
HisTyrValLeuCysValValThrIlePheSerValIleLeuLeuAlaIleValAlaLeu 208

FIG. 1A

U.S. Patent

Feb. 11, 2003

Sheet 2 of 13

US 6,518,414 B1

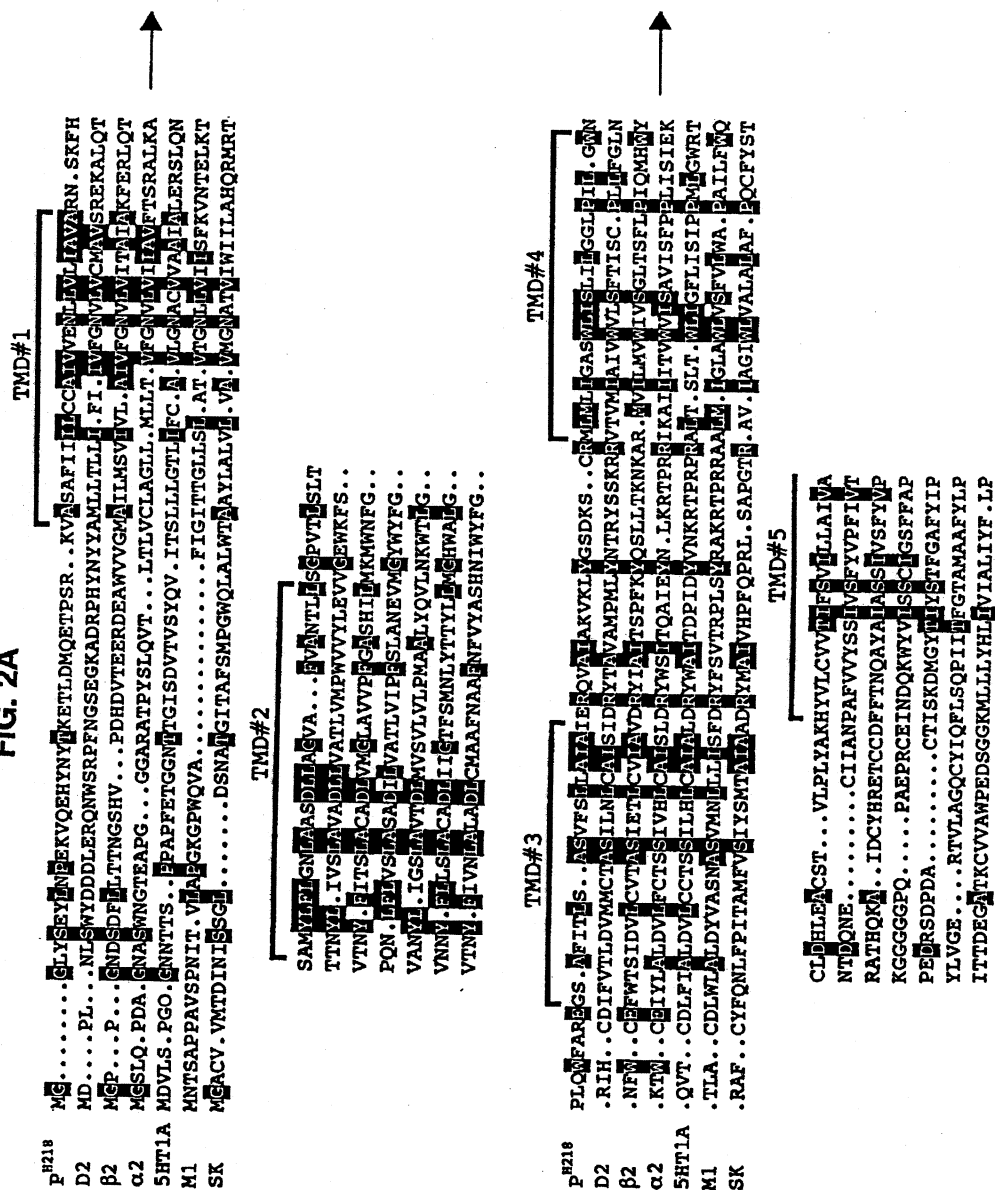
625	TACGTC CGAATCTACTTCGTAGTCCGCTCAAGCCATGCGGACGTTGCTGGTCTCCTCAGACG	228
	TyrValArgIleTyrPheValValArgSerHisAlaAspValAlaGlyProGlnThr	
685	CTGGCCCTGCTCAAGACAGTCACCATCTGCTACTGGGTGTTTTCATCATCTGCTGGCTGCCG	248
	LeuAlaLeuLeuLysThrValThrIleValLeuGlyValPheIleIleCysTrpLeuPro	
745	GCTTTTAGCATCCTTCTCTTAGACTCTACCTGTCCCGTCCGGCCGTGCTCCTGTCTCCTCTAC	268
	AlaPheSerIleLeuLeuLeuAspSerThrCysProValArgAlaCysProValLeuTyr	
805	AAAGCCCATTAATTCTTTCCTTCGCCACCCCTCAACTCTCTGCTCAACCCCTGTCATCTAT	288
	LysAlaHisTyrPhePheAlaPheAlaThrLeuAsnSerLeuLeuAsnProValIleTyr	
865	ACATGGCGTAGCCGGACCTTCGGAGGAGGTACTGAGGCCCTGCTGTGCTGGCGGCAG	308
	ThrTrpArgSerArgAspLeuArgArgGluValLeuArgProLeuLeuCysTrpArgGln	
925	GGGAAGGAGCAACAGGCGCAGAGGTGGGAACCCCTGGTCACCGACTCCTGCCCTCCGC	328
	GlyLysGlyAlaThrGlyArgArgGlyGlyAsnProGlyHisArgLeuLeuProLeuArg	
985	AGCTCCAGCTCCCTGGAGAGAGGCTTGCCATATGCCCTACATCGCCAAACATTTCTGGAGGGC	348
	SerSerSerSerLeuGluArgGlyLeuHisMetProThrSerProThrPheLeuGluGly	
1045	AACACAGTGGTCTGAGGGGAAATGTGAACCTGATCTGTAAACCAAGCCACAGAGAGCTCT	352
	AspThrValVal	

FIG. 1B

US 6,518,414 B1

FIG. 1C

FIG. 2A



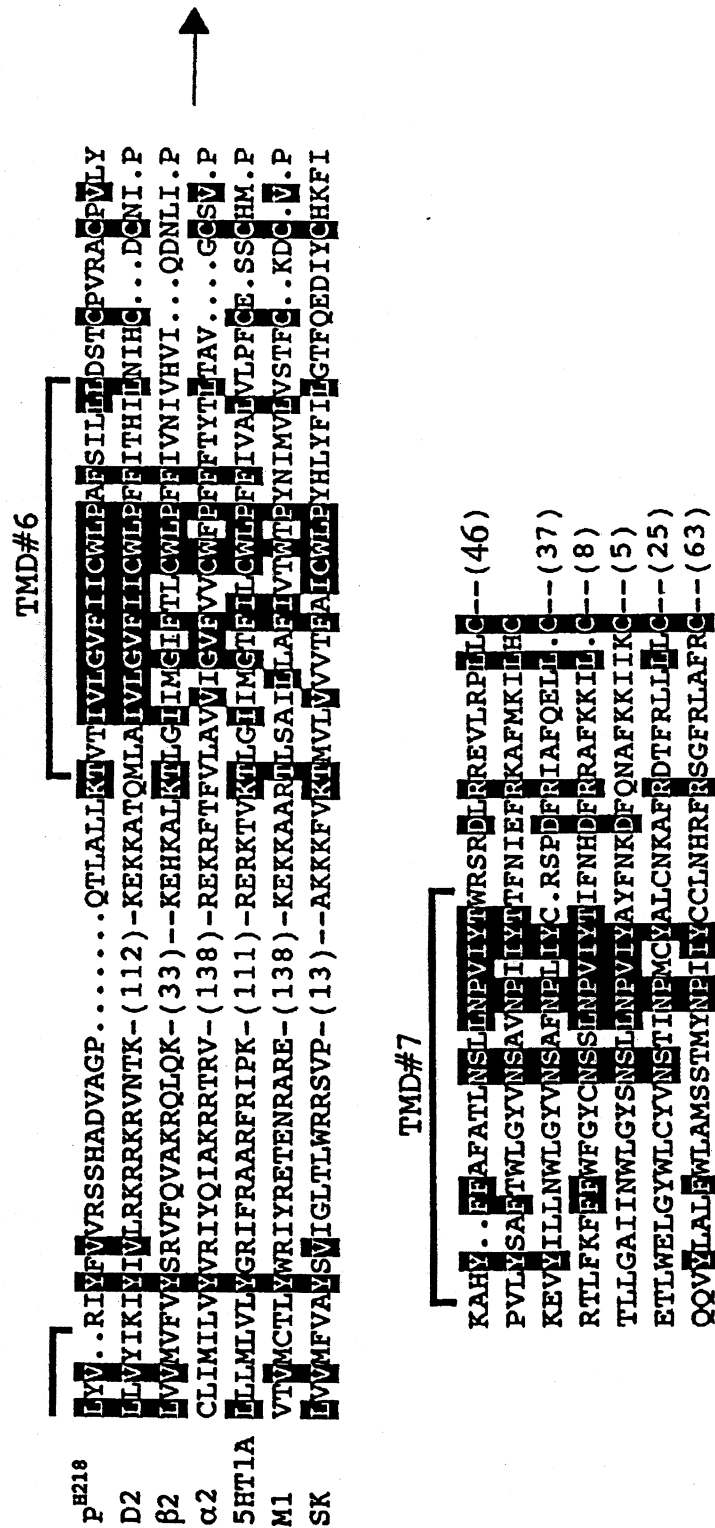
U.S. Patent

Feb. 11, 2003

Sheet 5 of 13

US 6,518,414 B1

FIG. 2B



U.S. Patent

Feb. 11, 2003

Sheet 6 of 13

US 6,518,414 B1

FIG. 3A

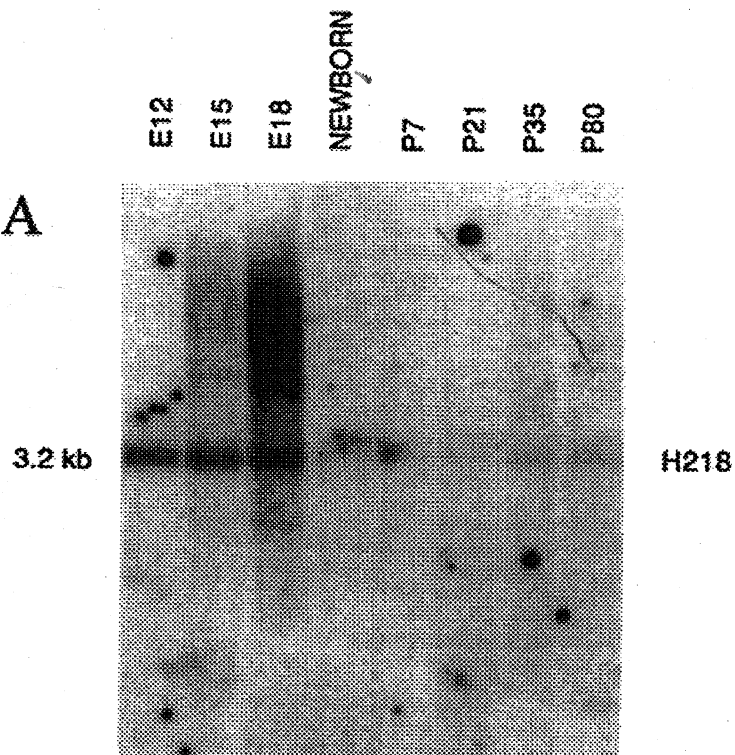
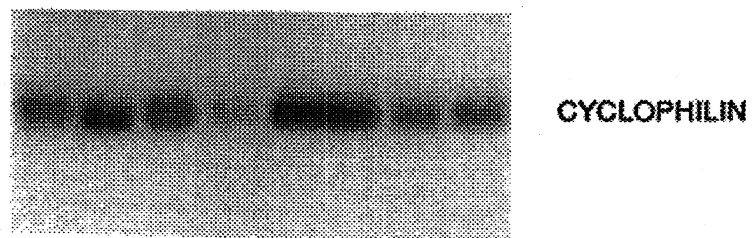


FIG. 3B



U.S. Patent

Feb. 11, 2003

Sheet 7 of 13

US 6,518,414 B1

FIG. 4A

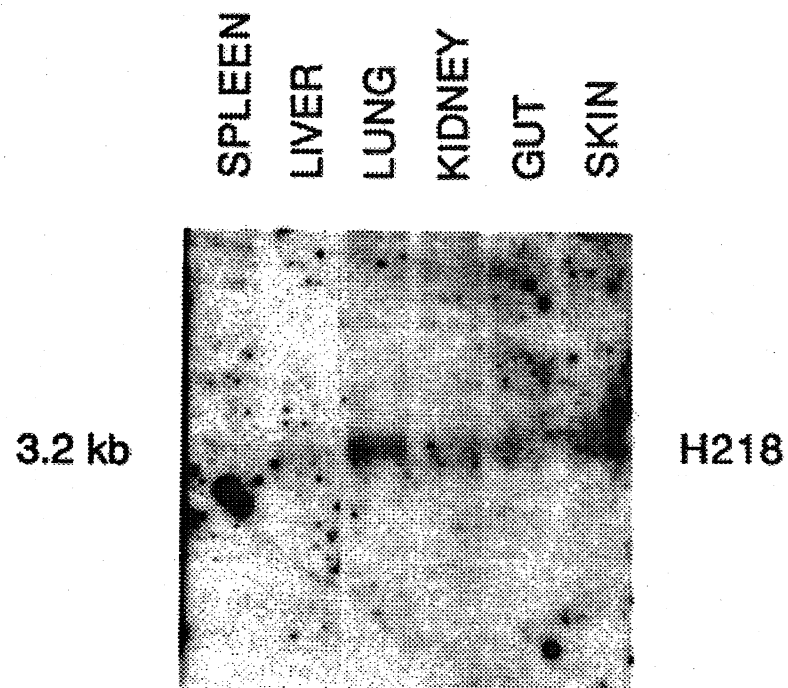
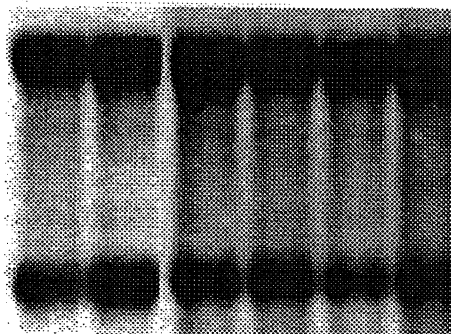


FIG. 4B



U.S. Patent

Feb. 11, 2003

Sheet 8 of 13

US 6,518,414 B1

FIG. 5A

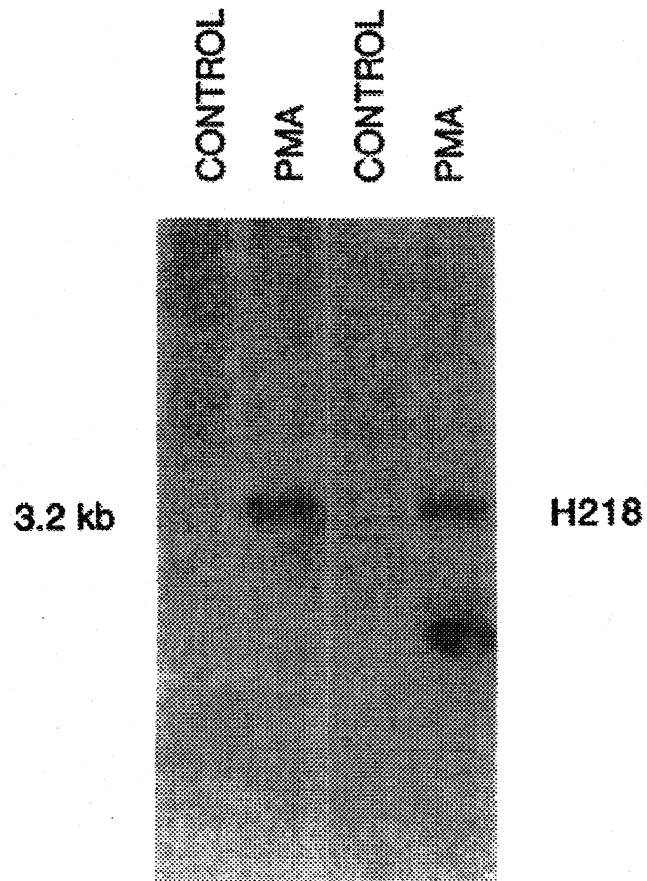
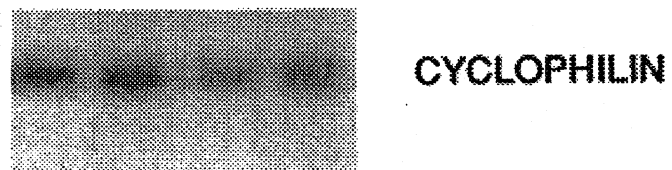


FIG. 5B



U.S. Patent

Feb. 11, 2003

Sheet 9 of 13

US 6,518,414 B1

FIG. 6A

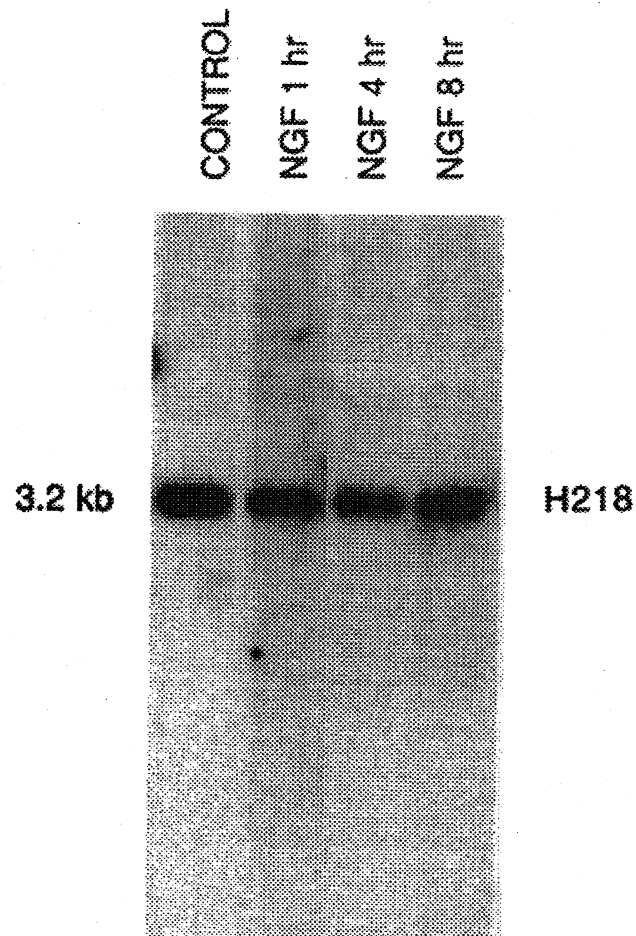
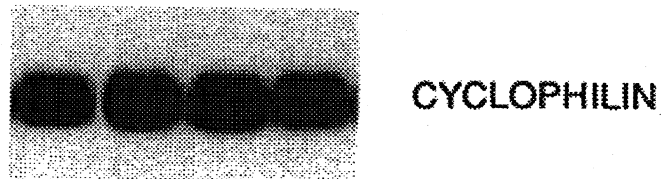


FIG. 6B



U.S. Patent

Feb. 11, 2003

Sheet 10 of 13

US 6,518,414 B1

-260 TTTGCTGGTCTCCGTCAGTCGCCGACAGCAAGATCGGGATCGCGGGTGTAG
-206 ACCCGGAGCCCCGGGACGAGCTTCGTCGCCGTTGAGCGAGGCTGTGTTCTCGGAGG
-146 CCTCTCCAGCCCAAGGAAAACATACATAAAAGCATCGGATTGCTGTGACCTGGCCTT
-86 GCTGTAACGTAAAGCTCGCTCAACCTCGCCCTCTAGCGTTTGTCTGGAGAAGTACCAACC
-26 CGGGCTCCTGGGACACAGTTGGGCTATGGTGTCTCCACAGCATCCAGTGGTTAAG
MetValSerSerThrSerIleProValValys 11

34 GCTCTCCGCAGCCAAAGTCTCCGACTATGGCAACTATGATATCATAGTCCGGCATTACAAC
AlaLeuArgSerGlnValSerAspTyrGlyAsnTyrAspIleValArgHisTyrAsn 31

94 TACACAGGCAAGCTGAACATCGGAGTGGAGAAGGACCATGGCATTAACCTGACCTTCAGTG
TyrThrGlyLysLeuAsnIleGlyValGluLysAspHisGlyIleLysLeuThrServal 51

154 GTGTTTCATTCTCATCTGCTGCTTGATCATCCCTAGAGAAATATTTGTCTTGCTAACTATT
ValPheIleLeuIleCysCysLeuIleIleLeuGluAsnIlePheValLeuThrIle 71

214 TGGAAAACCAAGAAGTTCACCGGCCCATGTACTATTTCATAGGCAACCTAGCCCTCTCG
TrpLysThrLysLysPheHisArgProMetTyrTyrPheIleGlyAsnLeuAlaLeuser 91

274 GACCTGTTAGCAGGAGTGGCTTACACAGCTAACCTGTGTCTGTGGGCCACCCACTAC
AspLeuLeuAlaGlyValAlaTyrThrAlaAsnLeuLeuLeuserGlyAlaThrThrTyr 111

334 AAGCTCACACCTGCCAGTGGTTTCTCGGGGAAGGAAGTATGTTTGTGGCTCTGTCTGCC
LysLeuThrProAlaGlnTrpPheLeuArgGluGlySerMetPheValAlaLeuSerAla 131

394 TCAGTCTTCAGCCTCCTGCTATCGCCATTGAGCGCTACATCACCATGCTGAAGATGAAA
SerValPheSerLeuLeuAlaIleAlaIleGluArgTyrIleThrMetLeuLysMetLys 151

454 CTACACAACGGCAGCAACAGCTCGCGCTCCTTCTGCTGATCAGTGGCCTGTGGGTGTCATC
LeuHisAsnGlySerAsnSerSerArgSerPheLeuLeuIleSerAlaCysTrpValIle 171

FIG. 7A

U.S. Patent

Feb. 11, 2003

Sheet 11 of 13

US 6,518,414 B1

514 TCCCTCATCCTGGTGGCTGCCCATCATGGCTGGAACATGCATCAGCTCGCTGTCACG
SerLeuIleLeuGlyGlyLeuProIleMetGlyTyrPasnCysIleSerSerLeuSerSer 191

594 TGCTCCACCGTCTCCCGCTCTACCAAGCACTATATCTCTTCTGCACCAACCGTCTTC
CysSerThrValLeuProLeuTyrHisLysHisTyrIleLeuPheCysThrThrValPhe 211

654 ACCCTGCTCCTGCTTCCATCGTCACTCCTCTACTGCAGGATCTACTCCTTGGTGAGGACT
ThrLeuLeuLeuSerIleValIleLeuTyrCysArgIleTyrSerLeuValArgThr 231

714 CGAAGCCGCGCTGACCTTCCGCAAGAACAATCTCCAGGCCAGCCGAGTCCGAGAAG
ArgSerArgArgLeuThrPheArgLysAsnIleSerLysAlaSerArgSerSerGluLys 251

774 TCTCTGGCCTTGCTGAAGACAGTGCATCTGCTGAGTGTCTTCTTCTTCTGCTGGGCC
SerLeuAlaLeuLysThrValIleIleValLeuSerValPheIleAlaCysTrpAla 271

834 CCTCTCTCATCCTACTACTTTTAGATCTGGGTGCCAGCGCAAGACCTGTGACATCCTG
ProLeuPheIleLeuLeuLeuAspValGlyCysLysAlaLysThrCysAspIleLeu 291

894 TACAAGCAGAGTACTTCTGTTCTGCTGCTGCTGAACCTCAGGTACCAACCCCATCATC
TyrLysAlaGluTyrPheLeuValLeuAlaValLeuAsnSerGlyThrAsnProIleIle 311

954 TACACTGTGACCAATAAGGACATGCCCGGCTTCTATCAGGATCATATCTTGTTCGCAA
TyrThrLeuThrAsnLysGluMetArgAlaPheIleArgIleIleSerCysCysLys 331

1114 TGCCCCAACGGAGACTCCGCTGGCAAATTCAGAGGCCCATCATCCCGGGCATGGAATTT
CysProAsnGlyAspSerAlaGlyLysPheLysArgProIleIleProGlyMetGluPhe 351

1194 AGCGCAGCAAAATCAGACAACTCCTCCCAACCCAGAGGATGATGGGACAATCCAGAG
SerArgSerLysSerAspAsnSerSerHisProGlnLysAspGlyAspAsnProGlu 371

1254 ACCATTATGCTCTCTGGAACGTCAATCTTCTTCTTAAACCGGAAGCTGTTGATACTG
ThrIleMetSerSerGlyAsnValAsnSerSerSer*** 383

FIG. 7B

U.S. Patent

Feb. 11, 2003

Sheet 12 of 13

US 6,518,414 B1

1314 TTGATTCTGGCTTCATCACTCACTACCCCTAGCATTTTCAAAAACATCTCTCTTCTCCACT
1374 GCTGCAAGGAAGAAGCAGCCGGAGCCTGAGAGAGGGAGGAAAGGAGAAATGTGCGGCTT
1434 GGTGATACCATGTTGTAGGTAGGTTATGATTATGAACAATGCCCTGGGAAGGTGGAGAT
1494 CAGATCTGCCCTGCAGAGGTTTCCTGCCCCCTCCTAATCTCTTCACTTCCCTTCAGTCGTT
1554 TCTGTTTATCCCCCATACTCTTTTCTTTTCTCCGTTTCTCTCATTTCCCTTCTCTACC
1614 ATCGCTTTCTTTTCTCTTTTAAATTTTGGGGCAACAAGGAATCCCAAAATGGA
1674 TATTGTGGAAAACATAGTGTGAATGACGGCAAGAATGGTGGTAAATCAAAAGATAAAT
1734 TAACTTCATAAGACTGCTATTCTGAATGCAACAATCTTGTACAGTCAGGACTGATAAAA
1794 TGGAGCAATCAGACATTTTCAGATGCCCGTCAATGTAAATCACCTACTTGACATTGTAT
1854 GCAATACATTCACACAAAAAAGCAAAATCTGTAGCCTTATTTTGAACAATACTGAACATCAT
1914 AAATACTCATGGTTTCACTCTGTCCAGCGCCCTAAGGACTATGCTGCTGTAATACAGGAA
1974 AACACAGCGGATGCCCTCCTATTAATAATGTCACTCAAGAAAAGTCTCTTGTAACGTAAA
2034 GGCAAAACACATGTAGCTACTGAGCTATGACTGTCCCTTGGTCACTCTATGGGAAAAACA
2094 CCGGACTCCAC

FIG. 7C

U.S. Patent

Feb. 11, 2003

Sheet 13 of 13

US 6,518,414 B1

5'-ccgcagacgctagccctgctcaagacggtcaccatcgtgctaggcgtc

P Q T L A L L K T V T I V L G V

tttatcgtctgctggctgcccgccttcagcatcctccttctggactatgcctgtcccgtc

F I V C W L P A F S I L L D Y A C P V

cactcctgcccgatcctctacaaagcccactacttttcgccgtctccaccctg -3'

H S C P I L Y K A H Y F F A V S T L

FIGURE 8

08CV3742

JUDGE PALLMEYER

MAGISTRATE JUDGE VALDEZ

TG

US 6,518,414 B1

1

MOLECULAR CLONING AND EXPRESSION OF G-PROTEIN COUPLED RECEPTORS

CROSS-REFERENCE TO A RELATED APPLICATION

This application is a continuation-in-part of application Ser. No. 08/760,936, filed on Dec. 6, 1996, now U.S. Pat. No. 5,856,443, which is a continuation of application Ser. No. 08/196,989, filed on Feb. 15, 1994, now U.S. Pat. No. 5,585,476.

This invention was made with government support under the National Institute on Drug Abuse grant number DA07244. The government has certain rights in this invention.

BACKGROUND OF THE INVENTION

The development of multicellular organisms requires the orchestration of many precisely coordinated events involving cell-type specific growth, proliferation, differentiation, migration, and cell death. Not surprisingly, intercellular communication plays critical roles in these processes. Although the molecular mechanisms involved in this communication are in general poorly understood, this research field is characterized by increasingly rapid progress initiated by the realization that viral oncogenes are, in many cases, transformed versions of cellular genes (proto-oncogenes) that participate in the intercellular communication directing development. Furthermore, it has been established that many non-viral forms of cancer also result from transformation of genes involved in signal transduction (e.g. growth factors, growth factor receptors, and transcription factors).

A large number of mammalian growth factor receptors have been cloned and many are recognized proto-oncogenes (Yarden and Ullrich, 1988). Most of these cloned receptors are members of a superfamily of integral membrane proteins with intrinsic, growth factor-inducible, tyrosine kinase activity. An extensive research literature now documents the critical roles these receptors play in cell proliferation, differentiation, and malignant transformation. However, multiple lines of evidence suggest that members of the G-protein coupled receptor (GPR) superfamily may also participate in mammalian development and oncogenesis. For example, both the yeast *S. cerevisiae* and the slime mold *D. discoideum* express GPRs that regulate cell differentiation (Devreotes, 1989; Sprague, 1991). In addition, mammalian mitogenesis and cell proliferation are affected by several peptides and neurotransmitters which are known to interact with GPRs (Hanley, 1989; Zachary et al., 1987).

Perhaps the most direct evidence linking GPRs with ontogeny and cancer has been provided by the ectopic expression of GPRs in tissue culture cells. Thus, the mas oncogene encodes a putative GPR (p^{mas}) and leads to malignant transformation when transfected into NIH3T3 mouse fibroblasts cells (Young et al., 1986). In addition, several serotonin and muscarinic acetylcholine receptors (all GPRs) also produce this malignant transformation if ectopically expressed in NIH3T3 cells and stimulated by their respective ligands (Gutkind et al., 1991; Julius et al., 1989; Julius et al., 1990). While these data illustrate that GPRs can greatly influence cell proliferation and morphology, the GPRs that were studied are unlikely to be involved in these processes in vivo because they reside in fully differentiated, postmitotic cells such as neurons where serotonergic receptors, muscarinic receptors, and most likely p^{mas} regulate the changing electrical properties of neuronal membranes involved in neurotransmission. However, these data

2

support the possibility that other GPRs are expressed in vivo in immature cells where they regulate proliferation and differentiation. Furthermore, these data suggest that some forms of cancer may result from mutations or viral infections that lead to improper functioning, activation, or expression of such GPRs. Thus, identification and characterization of such receptors should significantly advance both the study of normal development as well as the search for diagnostic and therapeutic tools in oncology.

BRIEF SUMMARY OF THE INVENTION

The subject invention concerns the cloning and sequencing of cDNAs and the proteins encoded by those cDNAs. The cDNAs encode novel polypeptides that are members of the G-protein coupled receptor (GPR) superfamily. The proteins encoded by the DNAs of the subject invention are involved in the regulation of cell proliferation and/or differentiation in vivo. The subject protein receptors are endogenously expressed in various tissues and cell lines.

Specifically, the subject invention concerns the cloning and sequencing of a rat cDNA (H218) that encodes a novel GPR designated p^{H218} . Further included in the subject invention are mammalian homologs, including the human homolog of the H218 cDNA. The H218 cDNA was used to determine that H218 mRNA is expressed in all developing organs tested and in seven out of seven cell lines tested. In addition, in the brain, H218 mRNA is much more highly expressed during a period of extensive proliferation and differentiation (embryogenesis) than a period of very limited cell proliferation and differentiation (adulthood), suggesting that p^{H218} does not function as a neurotransmitter receptor. Rather, p^{H218} functions as a growth factor ligand receptor.

The subject invention further concerns antibodies from animals immunized with peptides derived from p^{H218} GPR. Purified antibody made against one of the peptides recognizes a protein having an apparent molecular weight of 50-55 kDa as determined by Western blot analysis.

The subject invention also concerns cDNA of the rat-edg gene. Rat-edg cDNA encodes a GPR, $p^{rat-edg}$. The $p^{rat-edg}$ can be activated by some of the same ligand(s) that activate p^{H218} . By identifying compounds that specifically activate or inhibit this class of receptors one can develop unique, pharmaceutical therapies that effectively treat some forms of cancer.

A further aspect of the subject invention concerns polynucleotide molecules that are antisense to mRNA of H218 and rat-edg. The antisense polynucleotide molecules can be used to reduce or inhibit the expression of the subject protein by binding to the complementary mRNA transcripts.

The subject invention also concerns methods of use for the polynucleotide sequences, the encoded proteins, peptide fragments thereof, polynucleotide molecules that are antisense to the H218 and rat-edg sequences, and antibodies that bind to the proteins and peptides. Such use includes diagnostic and therapeutic applications of the subject invention.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGS. 1A-1C shows the nucleotide (SEQ ID NO.1) and deduced amino acid sequence (SEQ ID NO.2) of H218 cDNA. The sequence was compiled from that of "H2" cDNA (nucleotides 16 to 2505) and "18" cDNA (nucleotides-155 to 288) which are identical throughout the region of overlap. A black box highlights the optimal consensus sequence for translation initiation. A potential polyadenylation signal is double-underlined and a consensus

US 6,518,414 B1

3

sequence associated with mRNA instability is boxed. Repetitive nucleic acid sequences in the 3Q untranslated region are underlined. An arrow designates a predicted N-glycosylation site. A consensus sequence for proline directed kinases is underlined with a broken line. Brackets below the amino acid sequence indicate possible nucleotide binding site components in the carboxy-terminal and "third cytoplasmic loop" domains respectively.

FIGS. 2A and 2B shows a comparison of p^{H218} (SEQ ID NO.2) with other G-protein coupled receptors. Black boxes highlight residues identical to p^{H218} residues. D2=D2 dopaminergic receptor (SEQ ID NO.9); β 2= β 2 adrenergic receptor (SEQ ID NO.10); α 2= α 2 adrenergic receptor (SEQ ID NO.11); 5HT1A=1A serotonergic receptor (SEQ ID NO.12.); M1=M1 muscarinic receptor (SEQ ID NO.13); SK= substance K receptor (SEQ ID NO.14). The numbers in parentheses indicate the number of omitted residues.

FIGS. 3A and 3B shows an X-ray autoradiograph of a Northern blot illustrating the ontogenic regulation of H218 mRNA levels in the rat brain: Poly-A RNA was extracted from whole rat brain at embryonic days 12, 15, 18, Birth, postnatal days 7, 21, 35, and 80 (adult). The resulting blot was probed for H218 mRNA (panel A), stripped, and then probed with a cyclophilin cDNA (panel B) to control for variation in extraction, loading, and transfer (brain cyclophilin mRNA levels are reported to be stable from E12 to adult). The relative intensity of the cyclophilin bands have consistently paralleled results obtained from probing the same blots with an oligo-dT probe designed to hybridize to all mRNA poly-A tails.

FIGS. 4A and 4B shows an X-ray autoradiograph of a Northern blot illustrating the distribution of H218 mRNA in various tissues of the postnatal day 14 rat. Approximately 20 μ g of total RNA was loaded per lane. The blot was probed for H218 mRNA (panel A), stripped, and then probed for rat ribosomal RNA (panel B) as an extraction, loading, and transfer control.

FIGS. 5A and 5B shows an X-ray autoradiograph of a Northern blot illustrating the effect of PMA treatment on H218 mRNA levels in RJK88 fibroblasts. Poly-A RNA was extracted from 2 independent 100 mm plates of cells treated with PMA for 2 hrs (PMA) or 2 parallel plates of cells treated with vehicle (CONTROL). The resulting blot was probed for H218 mRNA (panel A), stripped, and then probed for cyclophilin mRNA (panel B) as an extraction, loading, and transfer control. Lanes are presented in pairs based on their relative mRNA content (as indicated by the cyclophilin data).

FIGS. 6A and 6B shows an X-ray autoradiograph of a Northern blot illustrating the effect of NGF treatment on H218 mRNA levels in PC12 cells. Poly-A RNA was extracted from 4 independent 100 mm plates of cells treated with NGF for either 1, 4, or 8 hrs or with a vehicle (CONTROL). The blot was probed for H218 mRNA (panel A), stripped, and then probed for cyclophilin mRNA (panel B) as an extraction, loading, and transfer control.

FIGS. 7A-7C shows the nucleotide (SEQ ID NO.3) and deduced amino acid sequence (SEQ ID NO.4) of rat-edg cDNA. An ATTTA motif is boxed in black.

FIG. 8 shows a partial nucleotide sequence (SEQ ID NO.15) of a cDNA that encodes a human p^{H218} polypeptide (SEQ ID NO.16).

BRIEF DESCRIPTION OF THE SEQUENCES

SEQ ID NO. 1 is the nucleotide sequence of the H218 cDNA.

4

SEQ ID NO. 2 is the deduced amino acid sequence of the p^{H218} protein encoded by the H218 cDNA.

SEQ ID NO. 3 is the nucleotide sequence of the rat-edg cDNA.

SEQ ID NO. 4 is the deduced amino acid sequence of the p^{rat-edg} protein encoded by the rat-edg cDNA.

SEQ ID NO. 5 is the amino acid sequence of a synthetic p^{H218} peptide designated peptide 1.

SEQ ID NO. 6 is the amino acid sequence of a synthetic p^{H218} peptide designated peptide 2.

SEQ ID NO. 7 is the amino acid sequence of a synthetic p^{H218} peptide designated peptide 3.

SEQ ID NO. 8 is the amino acid sequence of a synthetic p^{H218} peptide designated peptide 4.

SEQ ID NO. 9 is the amino acid sequence of a D2 dopaminergic receptor.

SEQ ID NO. 10 is the amino acid sequence of a β 2 adrenergic receptor.

SEQ ID NO. 11 is the amino acid sequence of a α 2 adrenergic receptor.

SEQ ID NO. 12 is the amino acid sequence of a 1A serotonergic receptor.

SEQ ID NO. 13 is the amino acid sequence of a M1 muscarinic receptor.

SEQ ID NO. 14 is the amino acid sequence of a substance K receptor.

SEQ ID NO. 15 is a partial nucleotide sequence encoding a human p^{H218} polypeptide.

SEQ ID NO. 16 is an amino acid sequence of a human p^{H218} polypeptide encoded by SEQ ID NO. 15.

DETAILED DISCLOSURE OF THE INVENTION

The subject invention concerns novel cDNAs (H218 and rat-edg) that encode G-protein coupled receptors. The proteins, designated p^{H218} and p^{rat-edg}, play important roles in cell proliferation and differentiation, and in disease states such as cancer.

It has been determined that the protein encoded by H218 polynucleotides is a receptor for sphingosine-1-phosphate (S1P). The research literature indicates that S1P can affect cellular processes potentially involved in many functions including nervous system development, nervous system responses to injury, tumorigenesis, metastasis, inflammation and heart function (Bunemann et al., 1996; Postma et al., 1996; van Koppen et al., 1996; Kawa et al., 1997; Yamamura et al., 1997). Therefore, agonists and antagonists for H218 may be of great clinical value in the treatment of disorders related to the above listed functions, and potentially other, as yet to be discovered, functions.

Rat H218 cDNA has been completely sequenced (SEQ ID NO. 1) and the amino acid sequence of the polypeptide that it encodes determined (SEQ ID NO. 2) (FIG. 1). The H218 cDNA contains a 1056 bp open reading frame that encodes a polypeptide of 352 amino acids. The 3Q untranslated region of H218 cDNA contains repetitive sequences, a consensus sequence for mRNA instability, and a series of terminal adenosines preceded by a potential polyadenylation site. The predicted cytoplasmic regions of p^{H218} contain potential nucleotide binding site components and a consensus sequence for proline directed kinases involved in cell division and growth factor responses.

Analysis of the deduced amino acid sequence of p^{H218} revealed that it is a member of the GPR superfamily (FIG. 2). Several features of p^{H218} are common to all other GPRs,

US 6,518,414 B1

5

including: 1) seven regions of hydrophobicity which are predicted to act as membrane spanning domains, 2) a consensus sequence for N-linked glycosylation in its predicted N-terminal extracellular domain, and 3) a conserved cysteine residue and several serine and threonine residues in its predicted intracellular C-terminal domain. In addition, p^{H218} contains many other residues which are highly conserved among most GPRs. However, p^{H218} is distinct from these GPRs in that it does not contain certain highly conserved residues. Perhaps most notable are the aspartate and tyrosine residues at the cytoplasmic end of the third transmembrane domain, and the cysteine residue at the extracellular end of the same transmembrane domain.

p^{H218} affects the course of cellular proliferation and/or differentiation events. Of all cloned proteins, p^{H218} is most homologous to human p^{edg} , a putative GPR implicated in endothelial cell differentiation. The possibility of a direct interaction between p^{H218} and growth-related intracellular proteins is suggested by the similarity between the predicted cytoplasmic region of p^{H218} and motifs of the src homology domain 2 (SH2) found in many cytoplasmic proteins that are critically involved in growth-related signal transduction, including several proteins encoded by oncogenes.

A further aspect of the subject invention concerns polynucleotide molecules which encode the human homolog of the rat H218 gene. Human cDNAs that hybridize with rat H218 cDNA were isolated from a human embryonic brain cDNA library. A cDNA sequence (SEQ ID NO.15) encoding part of the human H218 protein (SEQ ID NO.16) is shown in FIG. 8. Also contemplated within the scope of the present invention are human genomic H218 polynucleotide sequences, including polynucleotide sequences that flank the protein coding region. These sequences include regulatory sequences and intron sequences.

The human H218 protein is also contemplated within the scope of the invention. In one embodiment, the human H218 protein comprises the amino acid sequence shown in FIG. 8. Fragments and variants of the human H218 protein, including those that are biologically active or that are capable of ligand binding, are also within the scope of the invention.

The subject invention also concerns methods for screening for and identifying ligands of the H218 proteins of the invention. The H218 polynucleotides of the invention can be used to express the H218 protein in any of a wide variety of different expression systems. The H218 protein can then be used to identify H218 agonists and H218 antagonists. The agonists and antagonists can be identified based on their ability to bind to the H218 protein. For example, polynucleotides encoding H218 can be introduced into procaryotic or eucaryotic cells, thereby causing the cells to make H218 protein from the H218 polynucleotides. Therefore, these cells express more H218 protein than they would if the H218 polynucleotides had not been introduced into them. Consequently, H218 ligands can be identified by screening compounds for their ability to bind more to the cells that express greater levels of H218 protein. The ligands may be labeled with radioactive isotopes, or chemical modifications. Alternatively, unmodified ligands may be tracked with other approaches such as antibody recognition. In addition, ligands can also be identified based on their ability to activate H218 protein. When identifying H218 ligands based on their ability to activate H218, activation of H218 can be measured using any of a number of different methods known in the art. For example, one can measure changes in 1) H218-induced intracellular signal transduction events, 2) H218 conformation, 3) proteins bound to H218, and 4) H218-induced changes in cell behavior or morphology.

6

Ligands identified using the above methods are also within the scope of the invention.

H218-related nucleic acids can also be used to reduce expression of H218 protein in cells. The ability of ligands to bind to H218 protein or to activate H218, as discussed above, can then be used to identify H218 ligands by comparing cells which express native concentrations of H218 with cells in which H218 concentrations have been reduced. Methods to reduce H218 protein concentrations in cells include, for example, antisense techniques and homologous recombination techniques.

The subject invention also concerns H218-specific nucleic acid probes which can be used to identify mutations in the H218 gene. Identifying such mutations may be important in the scientific study of those diseases which involve mutations in the H218 gene. In addition, identifying the mutations may contribute to the clinical diagnosis, management and counseling related to these diseases. Similarly, antibodies raised against H218 protein sequence may be useful in identifying disease-related changes in H218 protein. The scientific and clinical uses of such antibodies would be the same as those outlined above for the H218-specific nucleic acid probes.

The subject invention also concerns methods for providing gene therapy to a patient in need of such therapy by introducing into the cells of the patient, by in vivo or ex vivo means, a polynucleotide vector that can increase or decrease H218 expression. A polynucleotide expression vector comprising a polynucleotide of the invention that encodes an H218 protein can be used to increase levels of H218 expression in cells. In one embodiment, the polynucleotide encodes the human H218 protein which comprises the amino acid sequence shown in FIG. 8 (seq id no.16). Similarly, H218 protein levels can be decreased in cells using, for example, antisense sequences or homologous recombination techniques. Methods of providing gene therapy are known in the art.

A further aspect of the subject invention concerns antibodies raised against synthetic peptides of p^{H218} . These peptides, designated as 1, 2, 3, and 4 (and corresponding to SEQ ID NO. 5, SEQ ID NO. 6, SEQ ID NO. 7, and SEQ ID NO. 8, respectively), correspond to separate extracellular and intracellular regions of p^{H218} . These peptides and their amino acid sequence are shown in Table 1.

TABLE 1

Amino Acid Sequences of p^{H218} peptides		
p^{H218} peptide		Sequence
peptide 1	SEQ ID NO. 5	KETLDMQETPSR
peptide 2	SEQ ID NO. 6	YSEYLNPEKVQE
peptide 3	SEQ ID NO. 7	RQGGATGRRGG
peptide 4	SEQ ID NO. 8	RSSSSLERGLHM

Polyclonal antibodies that react with the antigen peptides were raised in rabbits immunized with the respective peptide. Each antibody recognizes by an ELISA assay the specific peptide used as the immunogen. One of the antibodies, from a rabbit immunized with peptide 1 (SEQ ID NO. 5), was affinity purified and used in a Western blot with antigens from a cell line that expresses H218 mRNA. This antibody recognized a band of 50 to 55 kDa, and a band of 180 to 200 kDa in the Western blot. These antibodies can be used for detecting and purifying the p^{H218} protein through standard procedures known in the art. The antibodies can also be used for localization of p^{H218} in tissues using immunohistochemical techniques known in the art.

US 6,518,414 B1

7

The subject invention further contemplates the use of the protein and peptides to generate both polyclonal and monoclonal antibodies. Contemplated within the scope of the invention are antibodies to wild type forms of H218, as well as mutated forms of H218 protein. Monoclonal antibodies to p^{H218} , and peptide fragments thereof, can be produced using the teachings provided herein in combination with procedures that are well known in the art. Such antibodies can be produced in several host systems, including mouse, rat, and human. In one embodiment, the antibodies of the invention bind to human H218 protein.

Also included within the scope of the invention are binding fragments of the antibodies of the subject invention. Fab', F(ab')₂, and Fv fragments may be obtained by conventional techniques, such as proteolytic digestion of the antibodies by papain or pepsin, or through standard genetic engineering techniques using polynucleotide sequences that encode binding fragments of the antibodies of the subject invention.

A further aspect of the subject invention concerns the cloning and sequencing of the rat homolog of the human *edg* gene, which also encodes a GPR. This rat gene, designated *rat-edg*, is similar in sequence to the human *edg* gene. The *rat-edg* cDNA (SEQ ID NO. 3) encodes a protein, $p^{rat-edg}$ (SEQ ID NO. 4). The $p^{rat-edg}$ protein also has several features in common with other members of the GPR superfamily including 1) seven hydrophobic regions presumed to act as transmembrane domains, 2) a putative N-glycosylation site in the N-terminal domain, 3) putative phosphorylation sites in cytoplasmic domains, and 4) a conserved cysteine residue in the C-terminal domain.

The subject invention also concerns polynucleotide molecules having sequences that are antisense to mRNA transcripts of H218 and *rat-edg* polynucleotides. An administration of an antisense polynucleotide molecule can block the production of the protein encoded by H218 or *rat-edg*. The techniques for preparing antisense polynucleotide molecules, and administering such molecules are known in the art. For example, antisense polynucleotide molecules can be encapsulated into liposomes for fusion with cells.

As is well known in the art, the genetic code is redundant in that certain amino acids are coded for by more than one nucleotide triplet (codon). The subject invention includes those polynucleotide sequences which encode the same amino acids using a different codon from that specifically exemplified in the sequences herein. Such a polynucleotide sequence is referred to herein as an "equivalent" polynucleotide sequence. Thus, the scope of the subject invention includes not only the specific polynucleotide sequences depicted herein, but also all equivalent polynucleotide sequences encoding the polypeptides of the subject invention, and fragments or variants thereof.

The polynucleotide sequences of the subject invention can be prepared according to the teachings contained herein, or by synthesis of oligonucleotide fragments, for example by using a "gene machine" using procedures well known in the art.

The polypeptides of the subject invention can be prepared by expression of the cDNAs in a compatible host cell using an expression vector containing the polynucleotide sequences of the subject invention. The polypeptides can then be purified from the host cell using standard purification techniques that are well known in the art. Alternatively, the polypeptides of the subject invention can be chemically synthesized using solid phase peptide synthesis techniques known in the art.

8

The polypeptides of the subject invention can be used as molecular weight markers, as an immunogen for generating antibodies, and as an inert protein in certain assays. The polynucleotide molecules of the subject invention can be used as DNA molecular weight markers, as a chromosome marker, and as a marker for the gene on the chromosome.

The term "polynucleotide sequences" when used in reference to the subject invention can include all or a portion of the cDNA. Similarly, polynucleotide sequences of the subject invention also includes variants, including allelic variations or polymorphisms of the genes. The polynucleotide sequences of the invention may be composed of either RNA or DNA. More preferably, the polynucleotide sequences of the subject invention are composed of DNA.

As used herein, the term "isolated" means, in the case of polynucleotide sequences, that the sequence is no longer linked or associated with other polynucleotide sequences with which it would naturally occur. Thus, the claimed polynucleotide sequences can be inserted into a plasmid or other vector, to form a recombinant DNA cloning vector. The cloning vector may be of bacterial or viral origin. The vector may be designed for the expression of the polypeptide encoded by the polynucleotide sequence. The vector may be transformed or transfected or otherwise inserted into a host cell. The host cell may be either prokaryotic or eukaryotic, and would include bacteria, yeast, insect cells, and mammalian cells. For example, a bacterial host cell may be *E. coli*, and a mammalian host cell may be the PC12 cell line.

As used herein, the term "isolated" means, in the case of proteins, obtaining the protein in a form other than that which occurs in nature. This may be, for example, obtaining p^{H218} by purifying and recovering the protein from a host cell transformed to express the recombinant protein. In the case of antibodies, "isolated" refers to antibodies, which, through the hand of man, have been produced or removed from their natural setting. Thus, isolated antibodies of the subject invention would include antibodies raised as the result of purposeful administration of the proteins, or peptide fragments thereof, of the subject invention in an appropriate host.

The various genetic engineering methods employed herein are well known in the art, and are described in Sambrook, J., et al. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York. Thus, it is within the skill of those in the genetic engineering art to screen cDNA libraries, perform restriction enzyme digestions, electrophorese DNA fragments, tail and anneal vector and insert DNA, ligate DNA, transform or transfect host cells, prepare vector DNA, electrophorese proteins, sequence DNA, perform Northern, Southern and Western blotting, and perform PCR techniques.

Materials and Methods

Cloning of H218 cDNA. A "LAMBDA ZAP" cDNA library (Stratagene, La Jolla, Calif.) constructed using rat hippocampal RNA was screened at medium stringency with a 926 bp 5'Q EcoRI-Bgl II 3'Q fragment of a D2 dopamine receptor cDNA (MacLennan et al., 1990). The cDNA was labeled with ³²P by random hexamer priming. Nitrocellulose filters were incubated for 2 hrs at 42° C. in 5xSSPE (1xSSPE=0.15 M NaCl, 12 mM NaH₂PO₄H₂O, 1 mM EDTA, pH 7.4), 40% formamide, 0.15% SDS, 5xDenhardt's solution, 100 µg/ml denatured salmon sperm DNA, and 2 µg/ml polyadenylic acid. The filters were then incubated overnight in the same solution at 42° C. with the probe added (approximately 10⁶ cpm/ml). The